DOCUMENT RESUME

ED 059 834 RE 003 447

AUTHOR Bennett, Stan
TITLE An Evaluation of Some of Ashton-Warner's Assumptions About Beginning Reading.
PUB DATE 71
NOTE 51p.; Based on the author's doctoral dissertation, submitted to the University of Michigan, December, 1970

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Beginning Reading; Early Childhood Education; Emotional Response; Learning Characteristics; *Preschool Children; *Reading Research; *Retention; Word Frequency; *Word Recognition

ABSTRACT

A total of 14 four- and five-year-old girls learned to read two blocks of 12 words, each block consisting of (1) four words requested by each child (own words); (2) four words mentioned by Ashton-Warner as "one look" words for individual children (AW words); and (3) four words from the Scott-Foresman basal reader series (BR words). Measures of emotionality/arousal, meaningfulness, word frequency, and word length were obtained in the following ways: individual child's rating of word emotionality, percent Galvanic Skin Responses (GSR) deflection, Noble's meaningfulness (number of different free-associates by the child), Thorndike-Lorge estimate of word frequency, and number of letters in each word. No significant differences occurred between own, AW, and BR words on a 2-minute retention test; but 24-hour retention was greater for own words than for AW or BR words. Own words elicited significantly greater GSR's and were more meaningful than either AW or BR words. Children also rated own words as significantly more emotional than BR words. In the variables of word frequency and word length, own words significantly differed from BR words, but were similar to AW words. Findings and their implications were discussed. Figures, tables, a bibliography, and an appendix listing the own words are included. (AW)

An Evaluation of Some of Ashton-Warner's Assumptions About Beginning Reading*

Stan Bennett

Institute for Child Study
University of Maryland

ABSTRACT

A total of fourteen 4-5 year-old girls learned to read two blocks of 12 words, each block consisting of: 1) four words requested by each child (Own words); 2) four words mentioned by Sylvia Ashton-Warner (1963) as "one look" words for individual children (AW words); and 3) four words from the Scott-Foresman (1962) basal reader series (BR words). Although reading performance on the three types of words was similar during original learning or within two minutes thereafter, long-term retention was greater for Own words than for AW or BR words. Own words elicited significantly greater GSRs (Galvanic Skin Responses) and were more "meaningful" than either AW or BR words. Children also rated Own words as significantly more "emotional" than BR words. These data are seen as support for Ashton-Warner's assertion that Own words are more "meaningful" and more "emotional" than BR words and that Own words produce superior reading performance.

* This paper is based on the author's doctoral dissertation at the University of Michigan (December, 1970).

REVIEW OF THE LITERATURE

Introduction

In her word with Maori and New Zealand children, Sylvia Ashton-Warner (1963) has developed an approach to beginning reading which is essentially a variation on the language experience method. Ashton-Warner begins reading instruction by asking each child for words he would like to learn to read and write. These are one-word sentences for beginning readers, she insists. She writes the words, one by one, on stiff cards and gives them to the child to learn. Once the child has learned approximately 40 words (the "key vocabulary"), she assists him in writing his own stories and books. She also encourages children to read each other's books, and eventually, standard published materials. Ashton-Warner passionately explains the importance of this approach to reading:

> First words must have intense meaning for a child. They must be part of his being .... They must be made out of the stuff of the child itself. I reach a hand into the mind of the child, bring out a handful of the stuff I find there, and use that as our first working material. Whether it is good or bad stuff, violent or placid stuff, coloured or dun. To effect an unbroken beginning ..... In this dynamic material .... the Maori finds that words have intense meaning to him, from which cannot help but arise a love of reading. For it's here, right in this first word, that the love of reading is born, and the longer his reading is organic the stronger it becomes, until by the time he arrives at the books of the new culture, he receives them as another joy rather than as a labour. (p. 30-32)

Ashton-Warner found that key vocabulary words children asked for tended to be highly emotional words of "intense meaning," frequently centering around the emotions of "fear" (e.g., "spider," "fight," "alligator") and "sex" (e.g., "kiss," "darling," "together"). She feels that there is a world of difference between these words and the typical

basal reader fare ("come," "see," "ball," "jump"). What is perhaps most important, she observed that words children asked to learn to read were remembered better than basal reader words.

Ashton-Warner also concluded that the number of letters in a word had no relationship to the difficulty in learning to read the word. Several "backward" readers in the class, she noted, had trouble learning short, basal reader type words like "come" or "look," yet rapidly learned much longer words like "skellington" (sic), "frightened" or "together."

Emotional Components in Reading and Learning/Memory

The proposition that highly "emotional" words are learned more rapidly than those with neutral arousal properties has been tested by Olson and Pau (1966). They asked first and sixth graders to classify 46 nouns, verbs and adjectives of high Thorndike-Lorge (1944) word frequency as "Words I love," "Words I hate," and "Words that don't matter." "Love" or "hate" words were termed "emotional," whereas "words that don't matter were assigned to the "non-emotional" category. Three words rated by the child as "emotional" and three as "non-emotional" were subsequently selected for the learning trials.[1] In two separate experiments (one in which form class of the stimuli was controlled and one in which it was not) Olson and Pau found that first graders learned to read"emotional" words significantly more rapidly than "non-emotional" words.

---

[1]Experiment 1 "emotional" words: "gold," "burn," and "kill;" "non-emotional" words: "high," "many," "hall."

For indirect evidence on the importance of "emotion" in reading, we can turn to basic research on learning and memory. Weiner (1966) has recently reviewed evidence on the relationship between affect and recall. Although the studies in this area are replete with contradictory results, Weiner concluded that intensity of affect occuring at the time of original stimulus presentation is positively related to subsequent recall.

The relationship between arousal and memory (with arousal measured by means of the Galvanic Skin Response - GSR -) has been the subject of an interesting series of studies pioneered by Kleinsmith and Kaplan (1963, 1964). They distovered that after one presentation of the word stimuli, paired-associates eliciting high arousal showed inferior immediate recall when compared to pairs eliciting low arousal. However, when retention was measured 45 minutes or one day or one week after original learning, high arousal pairs were recalled more frequently than low arousal pairs. There was actually a noticeable increase in the recall of high arousal pairs compared to the number of correct responses occurring at the time of immediate recall (often called "reminiscence"). For low arousal pairs, performance showed the usual forgetting curve--that is, fewer correct responses after long-term delay than for immediate recall.

Two recent studies by Kaplan and Kaplan (1969, 1970) also demonstrated the presence of the reminiscence effect for high arousal pairs. However, forgetting occurred for both high and low arousal pairs by 18 minutes after original learning (but significantly less forgetting of high arousal pairs).

This reminiscence/retention effect has subsequently been reported by a number of other researchers under a variety of conditions (Levonian, 1967; Butter, 1970; Berlyne, Borsa, Craw, Gelman & Mandall, 1965; McLean, 1969; Batten, 1967; Howarth & Eysenck, 1968). Uehling & Sprinkle (1968) and Berlyne, Borsa, Hamacher & Koenig (1966) found no significant difference in immediate recall of words accompanied by white noise vs. no white noise presentation,[2] but superior long-term retention of items accompanied by white noise.

Meaningfulness/Word Frequency in Reading and Verbal Learning

Some definitions and procedures. Over the years the term "meaningfulness" has had a variety of connotations for educators and psychologists. Chall (1967) has noted that the basal reader programs have emphasized "reading for meaning." To assure that children have experienced words used in early primers, vocabularies of these materials have typically been limited to high frequency words in the language. Ashton-Warner has stressed that a child should read material having maximal concurrence with his experiences, but she relies on the unique interests and experiences of each child by helping him learn to read and write words that are important to each individual -- not words determined by tabulating the words that all children are likely to know and use.

Experimental psychologists have developed a number of procedures for measuring "meaningfulness." Noble (1952a) devised one of the most widely used techniques -- essentially a continuous free-association

---

[2]White noise appears to produce reductions in skin resistance, i.e., arousal (Berlyne & Lewis, 1963).

measure. Stimuli (words, nonsense syllables, etc.) are presented to a group of people who are asked to give as many free-associates to each stimulus as possible during a given period of time (e.g., 60 seconds per stimulus). The "meaningfulness" ("M") of each item is the average number of continuous associations given to a particular stimulus. "M" presumably measures the variety of past associations with a particular word--the greater the number of continuous free-associates, the greater the "meaningfulness" of the word.

Psychologists have typically measured the frequency of occurrence of a word in the language by means of the Thorndike-Lorge (1944) word count. Only low, positive correlations have been found between frequency of word usage and "M." Winnick and Kressel (1965) reported a product moment correlation of +.261 between Thorndike-Lorge word frequency and "M"; Paivio, Yuille & Madigan (1968) and Saltz (1967) both found a correlation of +.33.

Meaningfulness and word frequency in verbal learning. There appears to be a general tendency for "M" to facilitate both serial learning (Noble, 1952b; Braun & Heymann, 1958) and paired-associate learning (e.g., Kothurkar, 1963; Cieutat, 1961; Young, 1961; Martin, Cox & Boersman, 1965; Hunt, 1959; Hopkins & Schulz, 1969; and the review of Goss & Nodine, 1965). What little research there is on the relationship between "M" and long-term retention suggests that although high "M" does facilitate original learning, it has a slightly detrimental effect on long-term recall (Leeming, 1964; Young, Saegert, & Linsley, 1968).

High Thorndike-Lorge word frequency has been found to facilitate serial learning (Postman, 1961; Sumby, 1963), free recall (Hall, 1954;

Bousfield & Cohen, 1955; Murdock, 1960), and--under some conditions--paired-associate learning (Martin, 1964; Winnick & Kressel, 1965; Schwartz, 1965; Saltz & Modigliani, 1967; Hall, 1968; and Shapiro, 1969). However, Keppel (1968) has concluded from his extensive review of the literature that word frequency does not exert a significant impact on long-term retention.

Beginning reading and meaningfulness/word frequency. Athol Packer (1970) has recently compared the vocabularies of children from four cities around the country. All words were classified into 14 categories (e.g., fear, locomotion, animals, food, clothing, colors); then the percentage of Own and basal reader words falling into each of the 14 categories was calculated. Non-significant rank-order correlations between Own vs. basal reader vocabularies were the general rule, leading Packer to conclude that "...the children's own key vocabulary is more meaningful than the basal reader vocabulary" (p. 564). The term "meaningfulness" as used here is ambiguous, though apparently the author means the extent of congruence with the child's experiences and needs.

Wiley (1928) had 56 first-graders learn to read a total of 60 new words over a period of five weeks. In addition, he asked each child to free-associate to the words. He found a product moment correlation of +.55 between "quickness" of word associations and ease of learning to read these words. Words with "richer" associations (like "dinner"), he concluded, are easier for children to learn to read than words with few associates occurring only after latencies of more than three seconds (e.g., "they").

A number of investigators have evaluated the relationship between frequency of word usage and ease of learning to read words (Wiley, 1928;

Wheeler, 1938; Rickard, 1935). Wiley and Wheeler found non-significant correlations between word frequency and ease in learning to read words; Richard reported a low, positive correlation (+.39)--suggesting to Richard that word frequency has a significant but not "predominant" role in learning to read.

The Effect of Word Length on Beginning Reading

Gates and Boeker (1923) were two early people to investigate the effect of word length on beginning reading. They had kindergartners learn to read words of 3-10 letters in length. The results showed a tendency for longer words to be more difficult to learn than shorter words.

Rickard (1935) found that for the 30 words most easily recognized by the first, second and third graders in his study, the average number of letters per word was 3.25; for the 30 most difficult words, the mean was 4.25 letters per word. However, the first graders in the Wiley (1928) study made more errors on short words (2-4 letters) than long words (5-7 letters).

Some Hypotheses

Given the foregoing review of the literature, several predictions seem appropriate. First, Own words should elicit higher GSRs and be rated as more emotional than words supplied by someone other than the individual child (Others words). Also, significantly more continuous free-associates ("M") should be given to Own than to Others words. If greater GSR deflections do occur to Own words, one would expect some masking of the positive effect of these words when recall is measured

immediately after learning trials; 24 hours or several days later the superiority of Own words should emerge. Neither word length or word frequency should have a major impact on learning or retention of words children learn to read. The present study attempted to evaluate the appropriateness of these predictions.

METHODOLOGY

Subjects

A total of fourteen 4-5-year-old white, middle socioeconomic-class girls from a nursery school in Ann Arbor, Michigan, participated in this study. These children comprised the entire population of such girls in the school who agreed to participate in the study and who could not already read. Mean Peabody Picture Vocabulary TEST (PPVT) IQ score across all 14 children was 109.21 (SD = 7.81); mean chronological age was 61.71 months (SD = 6.90).

The school is a day-care facility designed primarily for children having only one parent residing in the home. Prior to the start of the study, children in one class had learned to read a few words, but during the time the experiment was in progress, no formal reading instruction was under way except for practice in recognizing each other's names.

Word Selection

Children learned to read two blocks of 12 words, each block consisting of: 1) four words requested by each child (Own words); 2) four words mentioned by Sylvia Ashton-Warner (1963) as "one look" words for individual children (AW words); and 3) four words from the Scott-Foresman (1962) New Basic Reading Program (BR words).

In selecting the AW and BR words (i.e., Others words), a pool of such words was first constructed for each category using words mentioned by Ashton-Warner in her book Teacher and from the vocabulary list of the Scott-Foresman pre-primers. Only words which can be used as nouns and/or verbs in connected discourse were included in the AW and BR word pools. Words in the AW/BR master word pools were also limited to those of three letters or more in length.[3] The common proper names used in the BR series (e.g., Dick, Sally) were not included in the word pool because of the contact children were experiencing with these words in their classrooms.

With the exception of four words,[4] a stratified random sampling procedure was used in selecting the actual AW and BR stimulus words-to-be-learned from the master word pools. One of the requirements for inclusion in the study was that both Block #1 and Block #2 AW and BR words contain two words which can be used as nouns and two as verbs.[5] Another stipulation was that where possible sound and semantic similarities within the AW and BR categories should be minimized (e.g., book and look were not placed in the same block, nor were kiss and darling). The following set of AW and BR words resulted from this selection process:

---

[3]Based on the experimenter's previous experience with "organic reading," it seemed likely that Own words would be of at least three letters in length. In order to make the three categories of words more comparable, the minimum number of letters for AW and BR words was set at three. As it turned out, only one of the Own words requested by the 14 children was a two-lettered word.

[4]Two AW words--"kiss" and "ghost"--and two BR words--"come" and "look"--were included because Ashton-Warner mentions them as examples of Own and basal reader words.

[5]Had the selection been completely random, AW words would have been predominantly nouns and BR words mostly verbs.

	AW words	BR words
Block 1	kiss	come
	thunder	ball
	spider	look
	hit	mother

	AW words	BR words
Block 2	sing	can
	ghost	get
	darling	did
	kill	book

The words remaining in the AW and BR word pools after completion of this selection process were assigned to: 1) a Filler Word cateogry; or 2) an auxiliary pool of AW and BR words to be used in case a child knew how to read one of the AW or BR words prior to the learning trials. Filler words were used in the long-term retention test at the conclusion of the study. The words included in the auxiliary AW and BR word pools were not needed, however, since no child could read any Block 1 or 2 AW/BR words prior to the beginning of the learning trials.

Own words were also limited to nouns and verbs, though because of the relatively small number of verbs requested and the difficulty which some children had in requesting Own words, it was not possible to restrict Own words to two nouns and two verbs per block of words. In each of the two blocks the first four words requested were used as the Own words, the remaining two being placed in a Filler Word category for use in the long-term retention test at the conclusion of Experiment II. (See Appendix A for a complete listing of Own words learned.)

Procedure

Children were seen individually in a quiet room of the school normally used as a storage area. The room was purposely made as barren as

possible. The child was seated at a long, low table--directly across from the "teacher." The experimenter sat behind a screen monitoring the GSR equipment on both Thursdays of the study. Other days the experimenter served as recorder of the child's responses, so as to facilitate interaction between "teacher" and child. The "teacher" was a graduate student in education who had no previous contact with these children. She was told nothing about the hypotheses being tested in this experiment.

Table 1 provides a calendar of events describing the various components of the experimental design. Experiment I (the data from Thursday and Friday of both weeks) involved a 2-minute and 24-hour retention test following only one study trial on each word. This study trial consisted of asking the child to trace each letter of the word with her finger while the teacher sounded out the word; then the child was instructed to say the word aloud. Two minutes after the "one look" each child was asked to read as many of the words as she could; then 24 hours later she was asked to try to read the words again. In both the 2-minute and 24-hour retention tests the words were presented to the child one-by-one for attempted recall. Experiment I (the "one look" portion of the study) was included in the design so as to evaluate Ashton-Warner's assertion that one-trial learning is more common for Own words.

Experiment II (which included the same words as Experiment I) consisted of alternating study trials (the child traced the letters of each word as the teacher sounded out the word--as in Experiment I) and test trials (the teacher laid down all 12 words on the table and the child pointed to any she knew and said them if she could). For the long-term

TABLE 1
CALENDAR OF EXPERIMENTAL EVENTS

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
				Week 1, Experiment I BLOCK #1 WORDS		
				a) 1 study trial b) Pegboard task c) 2-minute retention test	a) 24-hour retention test b) Continuous free-association ("M") c) Rated emotionality	
	Week 1, Experiment II BLOCK #1 WORDS			Week 2, Experiment I BLOCK #2 WORDS		
	a) Pre-test b) 3 study (S) trials alternating with 3 test (T) trials (S-T-S-T-S-T)	a) Pre-test b) 3 study (S) trials alternating with 3 test (T) trials (S-T-S-T-S-T)	a) Pre-test b) 3 study (S) trials alternating with 3 test (T) trials (S-T-S-T-S-T)	a) 1 study trial b) Pegboard task c) 2-minute retention test	a) 24-hour retention test b) Continuous free-association ("M") c) Rated emotionality	
	Week 2, Experiment II BLOCK #2 WORDS					
	a) Pre-test b) 3 study (S) trials alternating with 3 test (T) trials (S-T-S-T-S-T)	a) Pre-test b) 3 study (S) trials alternating with 3 test (T) trials (S-T-S-T-S-T)	a) Pre-test b) 3 study (S) trials alternating with 3 test (T) trials (S-T-S-T-S-T)			
	Week 3 Experiment II LONG-TERM RETENTION TEST Block 1 + Block 2 + Filler words (12 + 12 + 12 = 36 words					

retention test of Experiment II, the 24 words learned by each child during the experiment were randomly interspersed with the 12 Filler words. The aim of this second experiment was to determine whether differences in learning/retention of Own vs. AW/BR (Others) words would be present after a number of learning trials on the words.

Each word included in the study was written on a separate 5" x 8" lined card using large, lower-case letters.[6]

The specific events taking place each day were as follows:

Thursday, Week 1

a) Each child was asked for six words she would like to learn to read. Six AW and six BR words were also presented orally. For each word the child was asked to say the word and then sit quietly and think about it during a 15-second delay interval included so as to provide sufficient opportunity for a Galvanic Skin Response (GSR) deflection to occur.

These 18 words[7] were then shown to the child to assure that none of the words could already be read. (Four Own words could be read at this point; these were replaced by requesting more Own words from the child.) One study trial followed immediately thereafter on the 12 words of Block 1 (i.e., the 18 words minus the six Filler Words).

b) Next, the child was asked to place wooden pegs in a pegboard for two minutes.

---

[6]Though for three words--"Donald Duck," "Snow White," and "Casper"--the first letters were capitalized.

[7]Four Own words, 2 Own Filler Words; 4 AW words, 2 AW Filler Words; 4 BR words, 2 BR Filler Words.

c) Finally, the word cards were again presented one-by-one and the child asked to read the words (the 2-minute retention test). For all retention tests in the study the teacher only informed the child as to whether her response was correct or not; no prompting was provided.

Friday, Week 1

a) The twelve Block 1 words were presented and the child again was asked to read those words she could (the 24-hour retention test).

b) Children were next asked to give up to five free-associates to each of the 12 words so as to obtain an estimate of "M" (Noble, 1952a). Each child then classified the 12 words as "liked," "hated," or "don't matter" words (Olson & Pau, 1966). These two measures are more fully explained later in this section.

Monday/Tuesday/Wednesday, Week 1

a) A pre-test on the 12 words from Thursday/Friday of Week 1.

b) Three alternating study/test trials on these 12 words (see the earlier description of study/test trials).

Thursday/Friday, Week 2 --

Identical to Thursday/Friday, Week 1, except with a second set of 12 words (4 Own, 4 AW, 4 BR). Six words were again assigned to the Filler Word category.

Monday/Tuesday/Wednesday, Week 2 --

Same as Monday/Tuesday/Wednesday of Week 1, except with the second set of 12 words from Thursday/Friday, Week 2.

Monday, Week 3 -- Long-term retention test

Five days after completion of learning the second block of 12 words, and 12 days after completion of Block 1 learning, a long-term

retention test was given. This consisted of presenting one-by-one the 24 words in Blocks 1 and 2 plus the 12 Filler Words. Order of presentation of the 36 words was randomized using a table of random numbers. Midway through the presentation of the 36 words, a 1-minute rest period was provided.

Word Attributes

Measures of emotionality/arousal, "M", word frequency and word length were obtained in the following ways:

1) Individual child's rating of word emotionality

Each word to be learned was presented on Friday of each week and the child was asked to indicate whether the word was one she "liked," "hated," or "one that doesn't matter" (as described by Olson & Pau, 1966). If the child did not know what "doesn't matter" means, it was explained to her. "Liked/hated" words were coded as "emotional" and "don't matter" words as "non-emotional."

2) Percent GSR deflection

Skin resistance was recorded by means of the Kaplan and Fisher (1964) modifications of Lykken electrodes. The recording apparatus consisted of a wideband, constant current system designed by Kaplan and Hobart (1965). One electrode was fastened to the child's thumb and one to the palm. GSR recording took place only on Thursday of Weeks 1 and 2.

A 15-second period for GSR deflection was allowed in scoring the GSR tracings. GSR deflections up to five seconds prior to the utterance of an Own word were included since GSR deflections sometimes predated utterance of a requested word.

The percent GSR deflection used was the difference (in ohms) between the greatest resistance after the beginning of the 15-second period and the lowest skin resistance during that time span. If several deflections occurred, the difference between the greatest and least resistance was used. There were a number of cases where there was no deflection during the 15-second period, and some other cases where relaxation (not arousal) was occurring during the 15-second interval; these were scored as zero GSR deflections.

3) Noble's "meaningfulness" ("M")

The 12 words in each block were presented one-by-one on Friday of each week and the child was asked to indicate what word each made her think of. Five different free-associates per stimulus word were requested, or until the child indicated she could think of no more. In some instances, the child gave several words as a free-associate to the stimulus word; each of these phrases was treated as one free-associate. The number of different free-associates was used as the estimate of "M" for each word of each child (Possible range = 0 to 5).

4) Word frequency

The Thorndike-Lorge (1944) estimate of word frequency was used in this study. Words with Thorndike-Lorge ratings of AA (100 or more occurrences per million words) were coded as "3"; words with a rating of A (between 50 and 100 occurrences per million) were coded as "2", and words with 49 or fewer occurrences per million were given a code of "1". For the few Own

words (like "kleenex") not included in the Thorndike-Lorge count, the least frequent code (i.e., "1") was used.

5) Word length

The number of letters in each word was the score used. In the case of the five Own words which were actually two words in length, the space between the words was counted as one letter.

RESULTS

Experiment I[8]

Figure 1 portrays the relative recall of Own, AW and BR words 2 minutes and 24 hours after "one look" at the words. One-way repeated effects analyses of variance (see Table 2) showed no significant differences between Own, AW and BR words on the 2-minute retention test. However, 24 hours later significant differences were found between the three types of words. Tukey tests done to compare pairs of means indicated that only Own and BR words were significantly different on the 24-hour retention test ($p < .05$). As can be seen in Figure 1, no forgetting occurred for Own words over the interval from 2 minutes to 24 hours; both AW and BR words showed forgetting.

In an attempt to evaluate the relationship of these results to the arousal/retention effect, the Experiment I data were analyzed by collapsing across Own-AW-BR words and then dividing the words into those

---

[8]Three of the 14 children were not available for the 24-hour retention test of Experiment I; therefore the analysis of Experiment I data was limited to the 11 subjects completing both the 2-minute and 24-hour retention tests (N = 11).

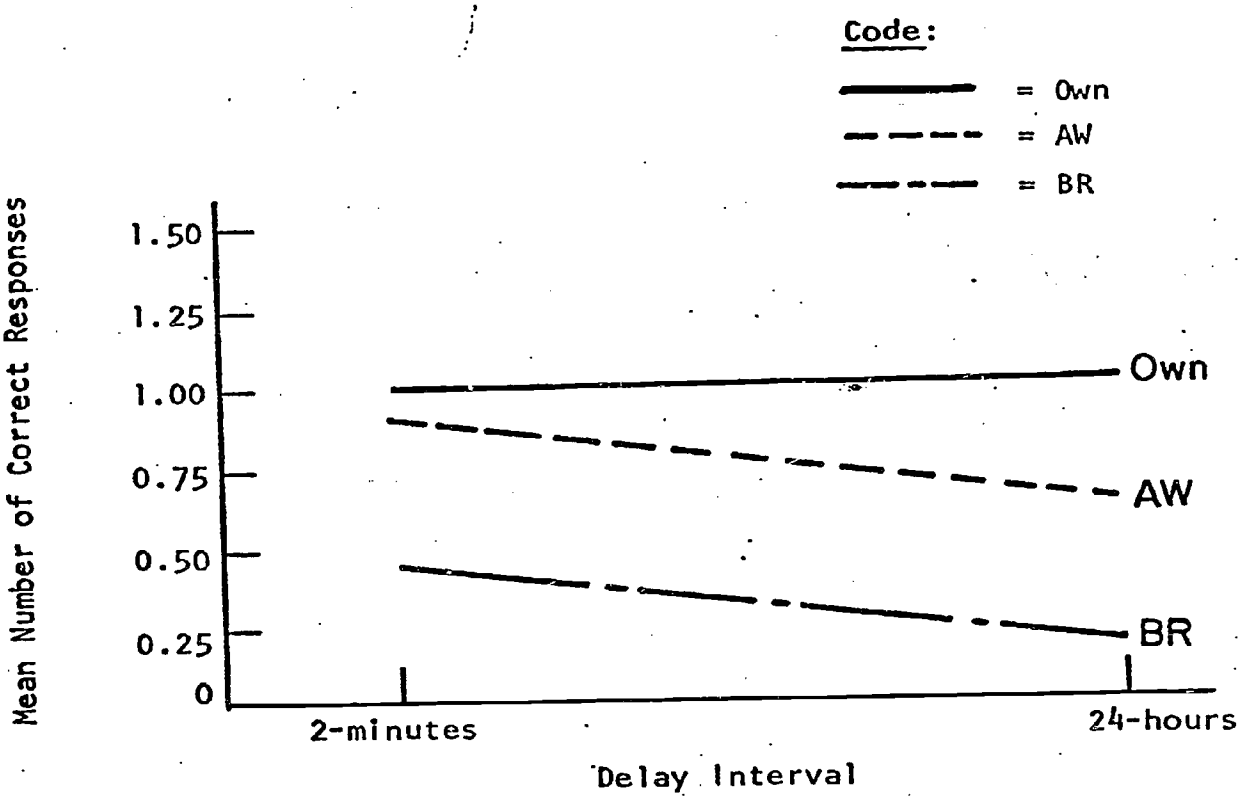

Figure 1: Mean recall of Own, AW, and BR words 2-minutes and 24-hours after "one look" (Experiment I).

TABLE 2

Summary of Analyses of Variance for 2-minute and 24 hour Retention Tests (Experiment I)

2-minute Retention

Source	SS	df	MS	F
Between Subjects	7.52	10		
Within Subjects	12.00	22		
Word type (O-AW-BR)	1.88	2	0.94	1.84 (N.S.)
Residual	10.12	20	0.51	

24-hour Retention

Source	SS	df	MS	F
Between Subjects	6.45	10		
Within Subjects	11.33	22		
Word type (O-AW-BR)	3.60	2	1.80	4.62**
Residual	7.73	20	0.39	

**Sig <.025 level

eliciting high (7% or greater) and low (6% or less) GSR deflections. Words producing high vs. low GSR deflections were not significantly different on the 2-minute retention test, though words yielding high GSR deflections were better recalled 24 hours after the "one look" ($F = 5.71$; 1, 10 df; $p < .05$).

Experiment II[9]

Learning/Retention of Own/AW/BR Words

A 2-way repeated effects analysis of variance was done to evaluate the significance of differences between the total number of correct responses for each word type and each day of Experiment II learning trials (collapsing across the three test trials each day--Monday, Tuesday, and Wednesday of Weeks 1 and 2). The results of this analysis (see Table 3) showed only a main effect of days; the main effect of word type (Own/AW/BR) was not significant, nor was the interaction between days and word type. Figure 2 depicts the similar performance occurring for Own-AW-BR words during learning trials of Experiment II.

Figure 3 represents mean Day 3 performance on the Own-AW-BR words of Experiment II compared with long-term retention of these three types of words.[10] A 2-way repeated effects analysis of variance on these data (see Table 4) indicated a significant main effect of retention interval and a significant interaction between word type (Own/AW/BR)

---

[9]One of the 14 children was not present for the long-term retention test of Experiment II and was thus excluded from the analysis of Experiment II data (N = 13).

[10]Mean Day 3 performance for Own, AW, and BR words was computed because the long-term retention test consisted of only one test trial.

TABLE 3

Summary of Analysis of Variance
for Days 1-2-3 and Word Type
(Experiment II)

Source	SS	df	MS	F
Total	4249.92	116		
Subjects	1640.14	12		
Word type (O-AW-BR)	21.15	2	10.57	<1
Day (1-2-3)	1530.53	2	756.26	75.32***
Word type x Day	3.78	4	0.94	<1
Error for word type	621.30	24	25.89	
Error for day	243.91	24	10.16	
Error for word type x day	189.11	48	3.94	

***Sig. < .01 level

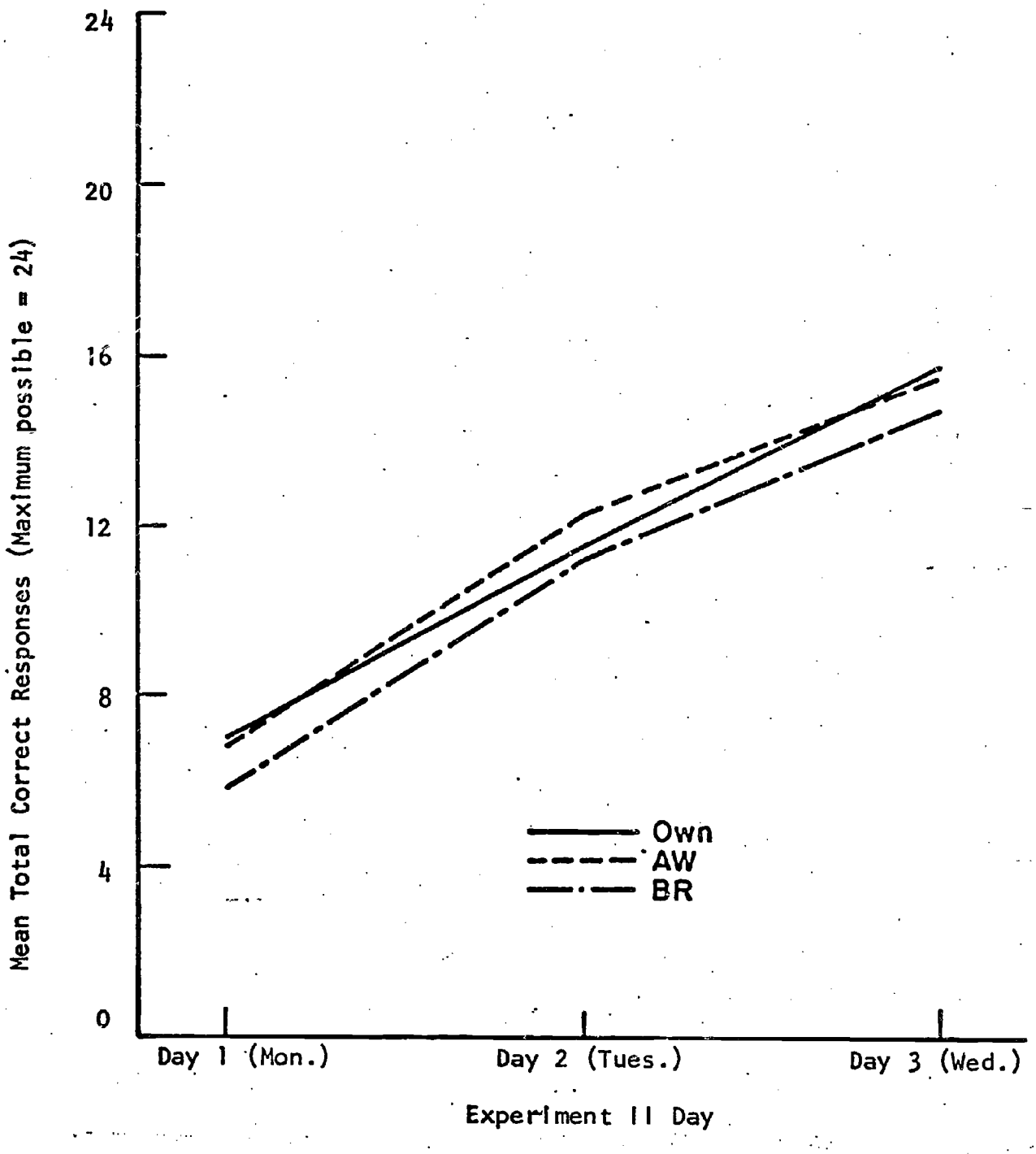

Figure 2: Mean total number of correct responses for Days 1, 2 and 3 (Monday/Tuesday/Wednesday) of Experiment II.

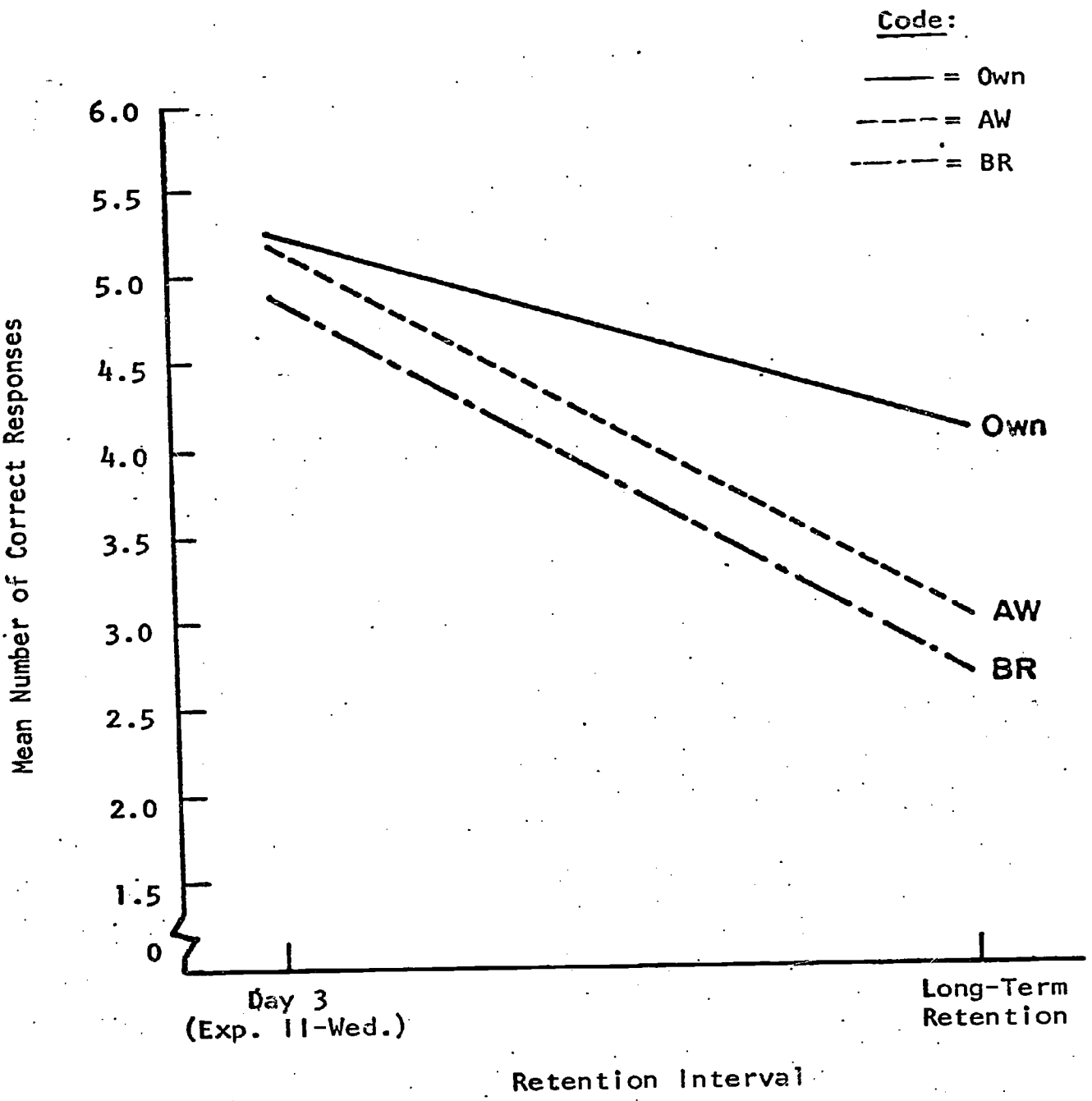

Figure 3: Relative forgetting of Own, AW, and BR words comparing Day 3 of Experiment II with Long-Term Retention

TABLE 4

Summary of Analysis of Variance for
Day 3 vs. Long-Term Retention

Source	SS	df	MS	F
Total	292.85	77		
Subjects	126.45	12		
Retention Interval (Day 3--LTR)	67.05	1	67.05	51.58***
Word type (O-AW-BR)	10.24	2	5.12	2.15
Word type x retention interval	4.44	2	2.22	4.44**
Error for retention interval	15.56	12	1.30	
Error for word type	57.13	24	2.38	
Error for retention interval x word type	11.98	24	0.50	

**Sig. < .025
***Sig. < .01

and retention interval. An analysis of the interaction effect (Bruning & Kintz, 1968, p. 120-122) revealed that the forgetting curve for Own words was significantly ($p < .025$) shallower than the curves for AW and BR words (Own vs. AW: $F = 6.48$, 1,24 df, $p < .025$; Own vs. BR: $F = 6.82$, 1,24 df, $p < .025$; AW vs. BR: $F = 0.02$, 1,24 df, N.S.). The AW and BR curves were remarkably parallel, as they were throughout the study. A one-way repeated effects analysis of variance done on the data from the long-term retention test indicated significant differences between word types ($F = 5.45$; 2,24 df; $p < .025$). Tukey tests on the significant effects of word type showed that Own words were significantly more often read than BR words ($p < .05$) and marginally better than AW words ($p < .058$). BR and AW words were read about equally often.

Attributes of Own vs. AW vs. BR Words

Mean "M", rated emotionality, GSR, word length and word frequency for Own, AW, and BR words are shown in Figure 4. One-way repeated effects analyses of variance were done in order to evaluate the significance of differences between means for "M", rated emotionality, and GSR. These analyses revealed that Own/AW/BR words were significantly different on "M", GSR and rated emotionality ("M": $F = 7.56$; 2,24 df; $p < .01$; GSR: $F = 9.81$; 2,24 df; $p < .01$; rated emotionality: $F = 5.33$; 2,24 df; $p < .025$). Tukey tests on the "M" and GSR attributes indicated that Own words were significantly higher than AW or BR words on both of these variables; AW and BR words were not significantly different on "M" or GSR. For rated emotionality, only Own and BR words were significantly different ($p < .01$).

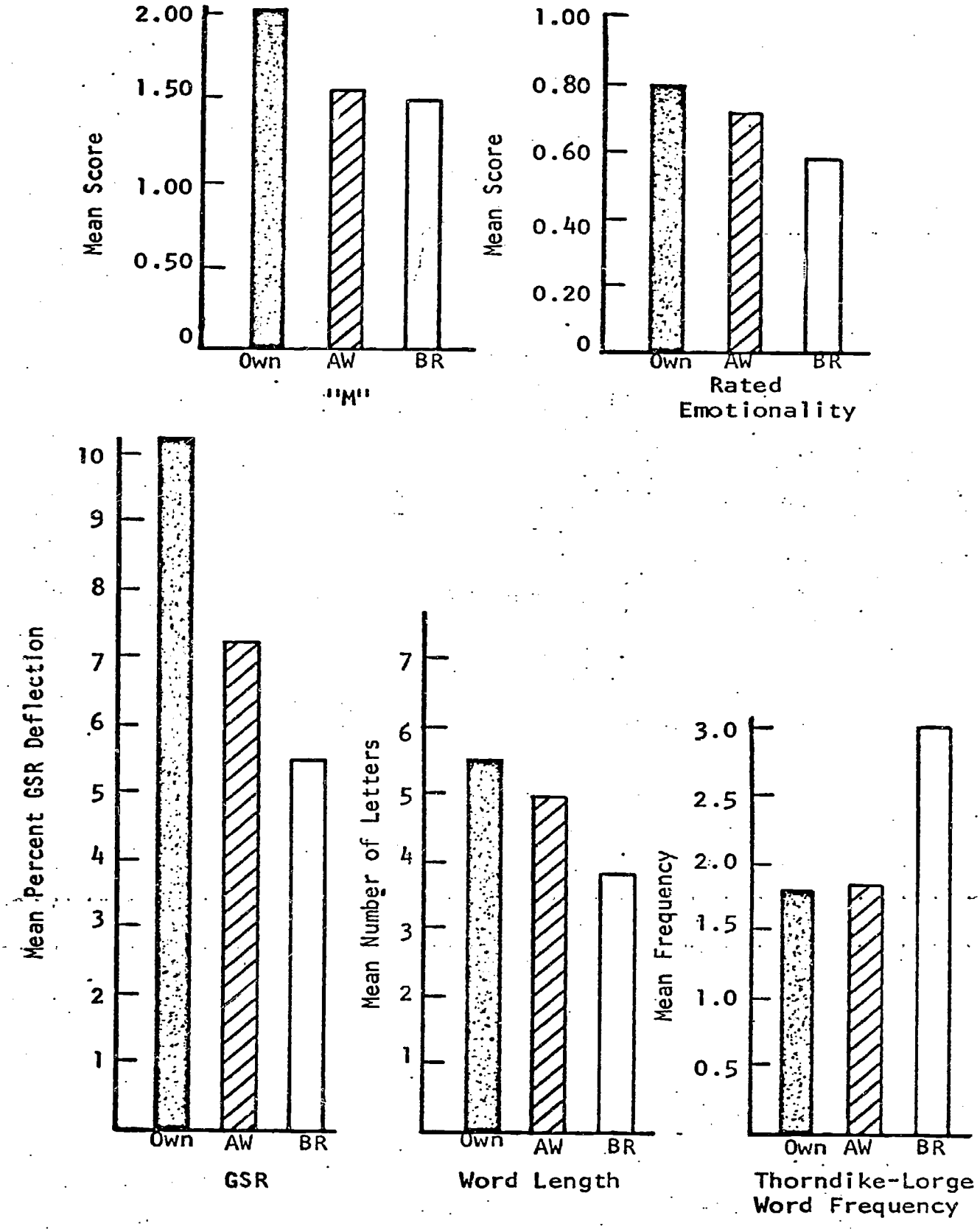

Figure 4: Mean scores on selected word attributes.

In the case of word length and frequency it was necessary to use the Median Test (Siegel, 1956) to determine if Own, AW and BR words were significantly different on these variables.[11] Own words were significantly longer than BR words ($p < .01$) but they were not longer than AW words; AW and BR words were not significantly different in word length. For Thorndike-Lorge word frequency, only Own words were significantly lower in frequency than BR words ($p < .05$).

Three fairly clear patterns emerged from these word attribute analyses:

1) Own words were high on "M"; rated emotionality; arousal (GSR deflections); longer in word length; and of low word frequency.
2) BR words were low on "M", rated emotionality, and GSR arousal; of short word length and very high word frequency.
3) AW words usually fell somewhere between Own and BR words on these attributes.

The Relationship Between Word Attributes and Performance Within Word Types

Two types of analyses--ipsative and normative--were used in an attempt to identify which of the word attributes contributed significantly to performance on Own, AW, and BR words. The normative analyses involved determining the median value for each word attribute across all 13 subjects of Experiment II; then a one-way analysis of variance was done for each word attribute to determine if total number of correct responses was significantly better for words high vs. low on that variable. For the ipsative analyses, each individual child's word attribute scores were divided at the median (or as close to that as possible).

---

[11] Non-homogeneity of variance for the three word types on the word length and word frequency dimensions precluded use of analysis of variance.

Mean performance for each child was then computed for words high vs. low on the various word attributes; finally, one-way analyses of variance were performed on these data.

Perhaps an example will help clarify these two types of analyses. The AW word "thunder" had a high "M" rating across all children and thus was placed in the high "M" normative category; however, for a particular child "thunder" might be low in "M" and thus be included in the low "M" ipsative category. Because word length and word frequency are the same for all AW and BR words (a constant set of AW and BR words was used), the ipsative and normative analyses are necessarily the same for these words. Normative analyses are not possible for Own words because each child received different words.

The performance measures used in these analyses were: 1) the mean total number of correct responses over the three days of the learning trials (Day 1+2+3); and 2) the mean total number of correct responses on the long-term retention test. An attribute-by-attribute discussion of the results of these analyses follows.

"M" ("Meaningfulness"). The ipsative analyses showed no significant effects for high vs. low "M" over the three word types. The normative analyses revealed that high "M" was associated with better Day 1+2+3 learning of: 1) BR words ($p < .005$); and 2) AW + BR words combined ($p < .05$). No differences in performance on the long-term retention test were noted for high vs. low "M" words on either the ipsative or normative analyses.

GSR. The ipsative data indicated no significant differences in either Day 1+2+3 learning or long-term recall for words eliciting high

vs. low GSR's. No significant effects werenoted in the normative analyses except that BR words yielding low GSR's were associated with better long-term retention ($p < .05$). However, it should be noted that GSR's were recorded during Experiment I; thus the generally non-significant effects of GSR for Experiment II are not especially surprising.

Rated emotionality. Because each Own/AW/BR word was either coded as "1" (word was liked/hated by the child) or "0" (word didn't matter to the child), ipsative analyses were not feasible since for each child typically only a couple of Own and AW words (at most) were rated as non-emotional. Under these conditions, dividing words into those high vs. low on emotionality would have produced spurious results, for in many cases performance on "low" emotionality words would have been based on two words or less.

Normative analyses for AW and BR words were done since they involved collapsing across rated emotionality for all words and all subjects. A normative analysis for AW + BR words (combined) would appear to be a close approximation to the Olson and Pau (1966) study. By thus collapsing across AW and BR words it was possible to assure a greater range of rated emotionality for a set of words determined by the experimenter. (Recall that Olson and Pau did not include an Own word category in their experiments.) The one-way repeated effects analyses of variance for AW + BR words on rated emotionality indicated that highly emotional words were more thoroughly learned than low emotional words during the learning trials of Experiment II ($p < .005$), and what's more, this facilitating effect of highly emotional words was also present on the long-term retention test ($p < .05$).

Word length and word frequency. As noted earlier, a normative word length/frequency analysis for Own words is not possible to do (subjects got different words); furthermore, ipsative and normative analyses for AW and BR words are identical (since all words were the same for all subjects). Consequently, for purposes of comparing performance on Own, AW, and BR words, the word length and word frequency analyses in this study will be considered ipsative.

Within the AW word category, short/high frequency words were learned more rapidly during the learning trials. However, for Own words, low frequency words were learned best; long and short words were learned equally well. No significant effect of long vs. short word length was noted for BR words during Day 1+2+3 learning. BR words were all of the highest coded Thorndike-Lorge word frequency and thus it was not possible to analyze the importance of this variable for BR words. On the long-term retention test, long Own words were recalled better than short ones; no other significant long-term retention differences occurred.

Qualitative Analyses for Specific Words

Performance was evaluated for specific words within the AW and BR word categories (see Table 5). Four words mentioned by Ashton-Warner as typical of "organic" vs. "inorganic" vocabularies ("kiss"/"ghost" vs. "come"/"look") will be discussed first. On the Day 1+2+3 learning trials and long-term retention test, the mean percent of these words recalled was:

		Days 1+2+3	Long-Term Retention
"Organic"	kiss	70.94%	69.23%
	ghost	31.62%	46.15%
"Inorganic"	look	47.01%	30.77%
	come	24.79%	0%

TABLE 5

Mean Rated Emotionality, GSR, "M", and Percent Correct Responses for Specific AW and BR Words (Experiment II)

	AW Words	Mean % Correct Responses Days 1+2+3	Mean % Correct Long-Term Ret.	Mean Emot.	Mean % GSR	Mean "M"
Block 1	kiss	70.94%	69.23%	0.62	8.62%	1.62
	thunder	35.94	23.08	0.69	8.85	1.92
	spider	47.01	23.08	0.92	8.15	1.46
	hit	46.15	15.56	0.69	4.31	1.38
Block 2	sing	53.85	38.62	0.46	1.92	1.31
	ghost	31.62	46.15	0.85	7.85	2.00
	darling	38.46	30.77	0.38	10.62	1.15
	kill	59.83	53.86	0.92	7.23	1.31
	BR Words					
Block 1	come	24.79%	0%	0.31	8.23%	1.38
	ball	48.72	30.77	0.85	5.92	1.62
	look	47.01	30.77	0.62	5.31	1.85
	mother	54.70	30.77	0.46	7.31	1.77
Block 2	can	63.25	69.23	0.38	4.23	1.23
	get	13.68	7.70	0.46	5.00	1.23
	did	31.62	38.62	0.38	5.46	0.85
	book	68.38	61.54	0.62	2.85	1.92

Thus, more correct responses occurred for "kiss" and "ghost" (combined) than for "look"/"come". However, if one were to examine performance on only those four words, a somewhat distorted picture would result. As may be seen in Table 5, children did rather poorly on learning/retention of some AW words--i.e., "darling", "thunder", "spider" (note that these are the longest AW words). There were some BR words that were learned and retained rather well--in particular, "can" and "book". In fact, of the first seven most easily learned words on Day 1-2-3 of Experiment II, four were BR words; on the long-term retention test three out of the first seven were BR words.

Table 6 was produced by collapsing over the AW and BR words and then dividing the resulting pool of words into those above and below the median on rated emotionality, GSR and "M". For rated emotionality and GSR, a greater number of AW words fell into the high emotionality group (66.7% of high emotional words were from the AW category; 75% of the high GSR words were AW words). An equal number of AW and BR words appeared in the high "M" category. There was a fair amount of overlap between high vs. low emotionality/BSR/"M". For example, "kiss", "thunder", "spider", and "ghost" were all high on emotionality/GSR/"M", whereas "sing", "can", "get", and "did" were all low on these variables.

Performance as a function of form class of the words. In order to try to determine which words were treated as nouns and which as verbs, each child's free-associate responses were consulted. Based on these responses, it was possible to decipher the probable form class which the child was thinking of when he heard the word. These data are presented in Table 7. It should be noted that "can" was consistently treated as

TABLE 6

AW and BR Words High and Low on Rated Emotionality, GSR, and "M"[a]

	High Emotionality Words (Decreasing Emot.)	Low Emotionality Words (Decreasing Emot.)
Rated Emotionality[b]	spider (AW) kill (AW) ghost (AW) ball (BR) hit (AW) thunder (AW) kiss (AW) look (BR) book (BR)	sing (AW) mother (BR) get (BR) darling (AW) can (BR) did (BR) come (BR)
	High GSR Words (Decreasing GSR)	**Low GSR Words (Descending GSR)**
GSR	darling (AW) thunder (AW) kiss (AW) come (BR) spider (AW) ghost (AW) mother (BR) kill (AW)	ball (BR) did (BR) look (BR) get (BR) hit (AW) can (BR) book (BR) sing (AW)
	High "M" Words (Decreasing "M")	**Low "M" Words (Decreasing "M")**
"M"	ghost (AW) thunder (AW) book (BR) look (BR) mother (BR) kiss (AW) ball (BR) spider (AW)	hit (AW) come (BR) sing (AW) kill (AW) can (BR) get (BR) darling (AW) did (BR)

[a]Rank ordered from highest to lowest scores within high and low rated emotionality, GSR, and "M".

[b]Because of tie scores, it was not possible to divide these words at the median. The 9 high emotionality words/7 low emotionality words division was the closest to an 8/8 break possible, given the data.

TABLE 7

Rank-ordering[a] of AW + BR Words for Days 1+2+3 and Long-Term Retention (Experiment II)

Days 1+2+3			Long-Term Retention		
AW/BR Words	Word Type	Probable Form Class	AW/BR Words	Word Type	Probable Form Class
kiss	(AW)	Verb	kiss	(AW)	Verb
book	(BR)	Noun	can	(BR)	Noun
can	(BR)	Noun	book	(BR)	Noun
kill	(AW)	Verb	kill	(AW)	Verb
mother	(BR)	Noun	ghost	(AW)	Noun
sing	(AW)	Verb	sing	(AW)	Verb
ball	(BR)	Noun	did	(BR)	Verb
look	(BR)	Verb	darling	(AW)	Noun
spider	(AW)	Noun	look	(BR)	Verb
hit	(AW)	Verb	mother	(BR)	Noun
darling	(AW)	Noun	ball	(BR)	Noun
thunder	(AW)	Noun/ Verb	thunder	(AW)	Noun/ Verb
ghost	(AW)	Noun	spider	(AW)	Noun
did	(BR)	Verb	hit	(AW)	Verb
come	(BR)	Verb	get	(BR)	Verb
get	(BR)	Verb	come	(BR)	Verb

[a]From highest percent correct to lowest percent correct.

a noun by children in this study (for example, as a free-associate to "can" one child said "tin"). The verbs resulting from this analysis were: "kiss", "kill", "sing", "did", "look", "hit", "get", and "come"; the nouns were "can", "book", "ghost", "darling", "mother", "ball", and "spider" ("thunder" seemed to be treated as both a noun and a verb). It can be seen in Table 7 that on Day 1-2-3 learning, the three best learned AW words were all verbs, while the best three BR words were nouns. The least well-read items were BR verbs, especially "get" and "come".

DISCUSSION

The Characteristics of Own, AW, and BR Words

The greater "emotionality" of Own words as compared with BR words was revealed in both the rated emotionality and GSR data. Though Own words were also accompanied by significantly greater GSR deflections than AW words, this was not true for rated emotionality. Of course, there is no reason why rated emotionality and GSR data should be identical. Rated emotionality, after all, was the child's subjective estimate of whether the word was "liked", "hated", or "didn't matter." GSR, on the other hand, was an estimate of the extent to which the child actually showed an arousal reaction to the stimulus word.

As Ashton-Warner predicted, Own words in this study were also more "meaningful" than Others words (as measured by "M"). Presumably, the process of generating an Own word requires the child to draw on past experiences and associations. Child #11 is a good example of this. She talked endlessly about her experiences with horses and frequently carried around a picture book depicting various sorts of horses. Two

of her Own words were "colt" and "horse." Another child (#10) had just been to the dentist; one of the words she requested was "teeth." Probably the more associations the child has had with the word, the more likely it will be retrieved from memory and requested as an Own word. Note that there is a very important difference between requesting an Own word and reading a word on a long-term retention test. Requesting an Own word involves retrieving a desired word from memory under conditions of minimal external stimulus control. Reading a word consists of giving an appropriate response when a specific stim is provided; the same may be said for paired-associate learning. "M" may not be a very important variable in long-term reading of words, but it appears to play a role in determining which Own words will be requested by children.

As noted in the review of the literature, basal readers have typically used high frequency words in hopes of assuring that the child has had experience with the word. BR words certainly do occur more frequently in the language than Own words, but the "M" data suggested that AW or BR words do not generally possess the richness of associations for <u>individual children</u> that Own words do.

For both word frequency and word length, Own words were significantly different from BR words; AW and Own words, however, were similar on these dimensions. The author suspects that the most likely explanation for this is that Own and AW words are a more representative sampling of words in the language than are BR words.

Learning and Retention of Own vs. Others Words: Some Possible Contributing Factors

One of the predictions made in this study was that if greater GSR deflections occurred to Own than to Others words the superiority of Own words would be seen only on a test of long-term retention. GSR deflections were significantly higher to Own words and the predicted pattern of results did occur.

Additional evidence that arousal as measured by GSR contributed to these results comes from several sources. In Experiment I high arousal items were not read more frequently on the 2-minute retention test, but they were on the 24-hour retention test. This finding is similar to the arousal/retention studies discussed earlier, where it was noted that the beneficial effect of high arousal items usually has appeared only on measures of long-term retention.

Secondly, the author knows of no research which indicates that "M" or word frequency or rated emotionality could have contributed to the pattern of results found in this study. (Own words, you will recall, were high on "M" and rated emotionality and of low word frequency.) As Keppel (1968) noted, word frequency does typically facilitate original learning, but it has no apparent effect on long-term retention. High "M" also facilitates original learning, but if anything it has a slightly detrimental effect on long-term recall (Leeming, 1964; Young, Saegert, & Linsley, 1968). The research of Olson and Pau (1966), in combination with evidence reviewed by Weiner (1966), suggests that both original learning and long-term recall should be facilitated by high rated emotionality. Experiment II of the present study corroborated these earlier findings. "M" and word frequency were in some cases related

to performance during original learning, but these positive effects disappeared on the long-term retention test. For AW+BR words, high emotionality was associated with both better original learning and long-term retention.

Thirdly, since Own words were read significantly more often than both AW and BR words on the long-term retention test of Experiment II, but AW and BR words were not different (Own > AW = BR), it would seem logical to look for word attributes showing an identical pattern if we want to explain the probable underlying processes. Only "M" and GSR show such a pattern. "M" can probably be eliminated as a contributor to the superior long-term reading of Own words for reasons mentioned earlier.

When the results of Experiments I and II are considered together with previous studies on meaningfulness, emotionality and arousal, a relatively strong case can be made for the importance of arousal in the present experiments.

There are, of course, some additional word attributes which must be considered as possible contributors to the superior long-term reading of Own words. Three such attributes will be discussed here: word length, form class of the words, and within-word-category associations.

It is conceivable that Own words (since they were longest of all words) were recognized by children through developing a strategy of identifying words by means of their length. But if word length can explain away the results of the Experiment II long-term retention test, why is it that AW and Own words (which were not significantly different in word length) were not recalled equally often? If word length were

that important, one would have expected that Own and AW words would have produced very similar long-term reading performance--and this was not the case.

A second way in which the effect of word length was evaluated was by collapsing across Own, AW, and BR words, then dividing the words into long and short categories (using an ipsative analysis). When a one-way repeated effects analysis of variance was done for Experiment II original learning and long-term retention, no significant differences in reading performance were obtained for long vs. short words. (Original learning: $F = <1$, 1,11 df, N.S.; Long-term retention: $F = 1.33$, 1,11 df, N.S.)

Another word attribute which might explain the long-term retention data is form class of Own vs. AW/BR words. As a quick glance at the Own, AW, and BR words will show, many more nouns were requested by children than occur in the AW and BR word categories. Could it be that the superiority of Own words is due to the greater frequency of nouns within the Own word category? Unfortunately, there is no airtight way to answer this, since only a few verbs were requested by children. What we can do is look at the relative long-term retention of nouns and verbs within the combined AW+BR word category (Experiment II). In Table 7 we note that of the seven words most frequently read on the long-term retention test, four were verbs and three were nouns; of the first five most often read words, three were nouns and two were verbs. However, AW verbs were retained better than AW nouns; the converse was true for BR words--BR nouns were more often remembered than were verbs. No very firm conclusion seems possible at this point, although in general verbs seem to be retained at least as well as nouns.

If one examines the list of Own words it is obvious that more semantic similarity is present within Own words than within AW and BR words. Apparently some children have a tendency to free-associate in the process of generating Own words (e.g., three Own words requested by child #11 were: "bluebird", "cardinal", and "robin"). Jenkins and Russell (1952) and Rosenberg (1965), among others, have shown that high association between words facilitates recall; therefore, the importance of this factor in the present study must be evaluated. This was done by having the experimenter and two graduate students independently list those word pairs in the Own word category which each rater judged as commonly associated with each other (e.g., "plate" and "fork" for subject #1). Next, Experiment II long-term retention scores were obtained for each child on both highly associated and non-associated words. Analyses of variance for all three judges revealed no tendency for highly associated Own words (e.g., "plate" and "fork") to be more frequently read than those judged as very infrequently associated (e.g., "airplane" and "cat").[12] It therefore appears that the degree of word association within Own words cannot account for the significantly superior long-tern reading of Own words.

The Importance of Word Frequency and Word Length Within Word Categories

It will be recalled that short, high frequency AW words were learned more rapidly than low frequency, long AW words. On the other hand, low frequency Own words were learned more easily and long words retained

---

[12]Judge #1: $F = 2.92$; 1, 102 df; N.S.; Judge #2: $F = 1.44$; 1,102 df; N.S.; Judge #3: $F < 1$; 1,102 df; N.S.

better on the long-term retention test of Experiment II. This seems to suggest that Ashton-Warner was accurate in her observation that children can read long Own words as well as short ones. However, she may not be correct in implying that long Others words are learned as rapidly as short ones. In fact, for original learning the date from the present study indicate that both Ashton-Warner and the basal reader people may be correct in their assumptions about which words are easier to learn. However, since the facilitating effect of short, high frequency AW words did not occur on the long-term retention test, the long-term importance of using short/high frequency words is probably minimal. For Own words, long word length is certainly no handicap in reading words several days after original learning.

Methods of Teaching Reading: Language Experience vs. Basal Readers

Dykstra (1968) reported in his summary of the now-famous cooperative research project in primary reading instruction that no differences in word recognition ability were found between language experience and basal reader programs. However, one very important flaw in these studies has been mentioned by Serwer (1969). She notes that the standardized tests used to evaluate word recognition skills contain a vocabulary bias favoring the basal reader program. As Serwer said:

> What existent instrument could have tested the wide range of idiosyncratic vocabulary elicited in the 24 first-grade Language Experience classrooms which were creating a corpus of stories about such varied experiences as turtles, custodial jobs, dinosaurs, shopping lists for cooking breakfast, etc? (p. 453)

In the present study, children were tested on precisely the same words they learned, thereby removing the bias of reading achievement tests.

Dykstra has noted that even within any one reading method included in the primary reading studies, there was such a wide range of performance for classrooms and projects that it was impossible to conclude that one method is invariably better than another. He suggested that perhaps evaluating a "method" is not the most fruitful approach to increasing the reading skills of children:

> It is likely that (reading) improvement would result from adopting certain elements from each of the approaches..... The first step would be to determine the elements within the various approaches most important to the success of that program. (p. 124)

The present experiment represented an initial attempt to look at some "elements" of Ashton-Warner's approach to beginning reading.

SOME IMPLICATIONS OF THIS STUDY

Having children read other children's "own" words will probably not produce much better reading performance than BR words. Support for this statement comes from the finding that the "own" words of Ashton-Warner's students (AW words) did not show much better long-term retention than BR words. However, data from this study suggest that if the teacher provides the words to be learned, words children rate as "emotional" will be learned and retained better than "non-emotional" words. Therefore, publishers of children's reading materials may improve reading performance somewhat by including words which children rate as "emotional." However, it must be emphasized that Own words still appear to produce the best long-term retention of words learned.

Many basal reader publishers have for years used short, high frequency words in their pre-primers and primers, with the apparent hope of facilitating the reading performance of children. In the present

study, short/high frequency AW words were in fact learned more rapidly than long/low frequency AW words during original learning, but these variables had no impact on long-term retention. There seems, therefore, to be no basis for the restriction of children's reading materials to such words.

One wonders how many teachers have been misled by reasonably high immediate recall of BR words, only to find on a reading achievement test at the end of the school year that children remembered few of these words. By the same token, it is conceivable that teachers who have tried using some variation on Ashton-Warner's reading approach have been unaware of the equivalent immediate recall to be expected for Own and BR words, but the superior long-term memory for Own words.

There is, of course, the important question of the extent to which children learn to break the grapheme-phoneme code. The present study really provides no data on this issue, but it may be that Own words are best for learning to break the code. That remains for future research to determine.

In the meantime, it appears that children are rather adept at choosing words which allow them to maximize their long-term reading performance.

BIBLIOGRAPHY

Ashton-Warner, Sylvia. Teacher. New York: Simon & Schuster, 1963.

Batten, D. E. Recall of paired associates as a function of arousal and recall interval. Perceptual & Motor Skills, 1967, 24, 1055-1058.

Berlyne, D. E., Borsa, D. M., Craw, M. A., Gelman, R. S., & Mandell, E. E. Effects of stimulus complexity and induced arousal on paired-associate learning. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 291-299.

Berlyne, D. E., Borsa, D. M., Hamacher, J. H., & Keonig, I. D. E. Paired-associate learning and the timing of arousal. Journal of Experimental Psychology, 1966, 72, 1-6.

Berlyne, D. E., & Lewis, J. L. Effects of heightened arousal on human exploratory behavior. Canadian Journal of Psychology, 1963. 17, 398-411.

Bond, G. L., & Dykstra, R. The cooperative research program in first-grade reading instruction. Reading Research Quarterly, 1967, 2, No. 4, 5-142.

Bousfield, W. A., & Cohen, B. H. The occurrence of clustering in the recall of randomly arranged words of different frequencies of usage. Journal of General Psychology, 1955, 52, 83-95.

Braun, H. W., & Heymann, S. P. Meaningfulness of material, distribution of practice, and serial-position curves. Journal of Experimental Psychology, 1958, 56, 146-150.

Bruning, J. L., & Kintz, B. L. Computational handbook of statistics. Glenview, Ill.: Scott-Foresman, 1968.

Butter, M. Johnna. Differential recall of paired associates as a function of arousal and concreteness-imagery levels. Journal of Experimental Psychology, 1970, 84, 252-256.

Chall, Jeanne. Learning to read: The great debate. New York: McGraw-Hill, 1967.

Cieutat, V. J. Group paired-associate learning: Stimulus vs. response meaningfulness. Perceptual & Motor Skills, 1961, 12, 327-330.

Dykstra, R. Summary of the second-grade phase of the Cooperative Research Program in primary reading instruction. Reading Research Quarterly, 1968, 4, No. 1, 49-70.

Fitzgerald, J. A. The vocabulary of children's letters written in life outside the school. Elementary School Journal, 1934, 34, 358-371.

Gates, A. I. A reading vocabulary for the primary grades. New York: Bureau of Publications, Teachers College, Columbia University, 1926, 1935, 1936.

Gates, A. I., & Becker, Eloise. A study of initial stages in reading by pre-school children. Teachers College Record, 1923, 24, 469-488.

Goss, A. E., & Nodine, C. F. Paired-associates learning. New York: Academic Press, 1965.

Hall, J. F. Learning as a function of word-frequency. American Journal of Psychology, 1954, 67, 138-140.

Hall, J. F. Paired-associate learning as related to "pronounceability" and word frequency. Canadian Journal of Psychology, 1968, 22, 212-216.

Hopkins, R. H., & Schultz, R. W. Meaningfulness in paired-associate recognition learning. Journal of Experimental Psychology, 1969, 79, 533-539.

Howarth, E., & Eysenek, H. J. Extraversion, arousal and paired-associate learning. Journal of Experimental Research in Personality, 1968, 3, 114-116.

Hunt, R. G. Meaningfulness and articulation of stimulus and response in paired-associate learning and recall. Journal of Experimental Psychology, 1959, 57, 262-267.

Jenkins, J. J., & Russell, W. A. Associative clustering during recall. Journal of Abnormal and Social Psychology, 1952, 47, 818-821.

Kaplan, S., & Fisher, G. R. A modified design for the Lykken zinc electrodes. Psychophysiology, 1964, 1, 88.

Kaplan, S., & Hobart, J. L. New Technique for recording skin resistance. American Journal of Medical Electronics, 1965, 4, 117-120.

Kaplan, S., & Kaplan, Rachel. Arousal and memory: A comment. Psychonomic Science, 1968, 10, 291-292.

Kaplan, Rachel, & Kaplan, S. The arousal-retention interval interaction revisited: The effects of some procedural changes. Psychonomic Science, 1969, 15, 84-85.

Kaplan, S., & Kaplan, Rachel. The interaction of arousal and retention interval: Ipsative vs. normative scoring. Psychonomic Science, 1970, 19, 115-117.

Keppel, G. Retroactive and proactive inhibition. In T. R. Dixon & D. L. Horton (Eds.) Verbal Behavior and General Behavior Theory. Englewood Cliffs, N.J.: Prentice-Hall, 1968.

Kleinsmith, L. J. & Kaplan, S. Paired-associate learning as a function of arousal and interpolated interval. Journal of Experimental Psychology, 1963, 65, 190-193.

Kleinsmith, L. J., & Kaplan, S. Interaction of arousal and recall interval in nonsense syllable paired-associate learning. Journal of Experimental Psychology, 1964, 67, 124-126.

Kothurkar, V. K. Effect of stimulus-response meaningfulness on paired associate learning and retention. Journal of Experimental Psychology, 1963, 65, 305-308.

Leeming, F. C. Effects of stimulus and response meaningfulness on the long-term retention of a paired-associate verbal learning task. Dissertation Abstracts, 1964, 25, 648.

Levonian, E. Retention of information in relation to arousal during continuously-presented material. American Educational Research Journal, 1967, 4, 103-116.

Martin, J. G. Associative strength and word frequency in paired-associate learning. Journal of Verbal Learning and Verbal Behavior, 1964, 3, 317-320.

Martin, C. J., Cox, D. L., & Boersma, F. J. The role of associative strategies in the acquisition of P-A material: An alternate approach to meaningfulness. Psychonomic Science, 1965, 3, 463-464.

McLean, P. D. Induced arousal and time of recall as determinants of paired associate recall. British Journal of Psychology, 1969, 60, 57-62.

Murdock, B. B., Jr. The immediate retention of unrelated words. Journal of Experimental Psychology, 1960, 60, 222-234.

Noble, C. E. An analysis of meaning. Psychological Review, 1952, 59, 421-420. (a)

Noble, E. C. The role of stimulus meaning (m) in serial verbal learning. Journal of Experimental Psychology, 1952, 43, 437-446. (b)

Olson, D. R., & Pau, A. S. Emotionally loaded words and the acquisition of a sight vocabulary. Journal of Educational Psychology, 1966, 57, 174-178.

Packer, Athol. Ashton-Warner's key vocabulary for the disadvantaged. Reading Teacher, 1970, 23, 559-564.

Paivio, A., Yuille, J. C., & Madigan, S. A. Concreteness, imagery, and meaningfulness values for 925 nouns. Journal of Experimental Psychology, 1968, 76, No. 1, Pt. 2, 1-25.

Peabody Picture Vocabulary Test. Minneapolis, Minn.: American Guidance Service, 1959.

Postman, L. Extra-experimental interference and the retention of words. Journal of Experimental Psychology, 1961, 61, 97-110.

Rickard, G. E. The recognition vocabulary of primary pupils. Journal of Educational Research, 1935, 29, 281-292.

Rosenberg, S. Association and phrase structure in sentence recall: A replication. In J. C. Ca' ord (Ed.) Studies in Language and Language Behavior, VI. Ann rbor, Michigan: Center for Research on Language and Language Behavior, February 1, 1968.

Saltz, E. Thorndike-Lorge frequency and m as separate factors in paired-associates learning. Journal of Experimental Psychology, 1967, 73, 473-478.

Saltz, E., & Modigliani, V. Response meaningfulness in paired associates: T-L frequency, m, and number of meanings (dm). Journal of Experimental Psychology, 1967, 75, 313-320.

Schwartz, F. A joint investigation of associative and word frequency in paired associate learning. Psychological Reports, 1965, 16, 1165-1172.

Scott-Foresman New Basic Readers. (Helen Robinson, Marion Monroe, & A. S. Artley, Eds.) Chicago: Scott, Foresman & Co., 1962.

Serwer, Blanche L. Linguistic support for a method of teaching beginning reading to black children. Reading Research Quarterly, 1969, 4, No. 4, 449-467.

Shapiro, S. I. Response word frequency in paired-associates learning. Psychonomic Science, 1969, 16, 308-309.

Siegel, L., & Seigel, Lila C. The instructional gestalt. In L. Siegel (Ed.) Instruction: Some Contemporary Viewpoints. San Francisco: Chandler Publishing Co., 1967, pp. 261-290.

Siegel, S. Non-parametric statistics for the behavioral sciences. New York: McGraw-Hill, 1956.

Sumby, W. H. Word frequency and serial position effects. Journal of Verbal Learning and Verbal Behavior, 1963, 1, 443-450.

Thorndike, E. L., & Lorge, I. The teacher's word book of 30,000 words. New York: Bureau of Publications, Teachers College, Columbia University, 1944.

Uehling, Barbara S., & Sprinkle, R. Recall of a serial list as a function of arousal and retention interval. Journal of Experimental Psychology, 1968, 78, 103-106.

Weiner, B. Effects of motivation on the availability and retrieval of memory traces. Psychological Bulletin, 1966, 65, 24-37.

Wheeler, L. R. A study of the relative difficulty of a primary reading vocabulary. Pedagogical Seminary and Journal of Genetic Psychology, 1938, 52, 183-201.

Wiley, W. E. Difficult words and the beginner. Journal of Educational Research, 1928, 17, 278-289.

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1962.

Winnick, Wilma A., & Kressel, K. Tachistoscopic recognition thresholds, paired-associate learning, and free recall as a function of abstractness-concreteness and word frequency. Journal of Experimental Psychology, 1965, 70, 163-168.

Young, R. K. Paired-associate learning when the same items occur as stimuli and responses. Journal of Experimental Psychology, 1961, 61, 315-318.

Young, R. K., Saegert, J., & Linsley, D. Retention as a function of meaningfulness. Journal of Experimental Psychology, 1968, 78, 89-94.

APPENDIX A

Own Words

	Block 1	Block 2
Child #1	fork	friend
	cat	elephant
	airplane	color
	plate	climb
Child #2	stairs	swing
	clock	home
	watch	supper
	heater	spy
Child #3	tree	color
	kleenex	climb
	leaf	games
	glasses	water
Child #4	elevator	fawn
	banister	dog
	slide	cub
	Snow White	restaurant
Child #5	cowboy	dollie
	guns	tricycle
	turtle	bear
	giraffe	climb
Child #6	corn	rope
	pen	color
	no	friends
	swing	kick
Child #7	paper	bear
	witch	dam
	pencil	games
	bat	sandbox

	Block 1	Block 2
Child #8	lipstick	bag
	ink	park
	eskimo	glasses
	kleenex	names
Child #9	paper clip	tack
	railing	glass
	window	wood
	handle	dirt
Child #10	plane	pegs
	teeth	board
	magic marker	radiator
	leaf	screen
Child #11	bluebird	baby
	cardinal	eat
	robin	climb
	Casper	tree
Child #12	leaf	bee
	table	turkey
	bracelet	color
	kleenex	Donald Duck
Child #13	shirt	teddy bear
	ring	outside
	necklace	grass
	pencil	friends
Child #14	colt	snore
	horse	toyshop
	glue	kitty
	ceiling	candy